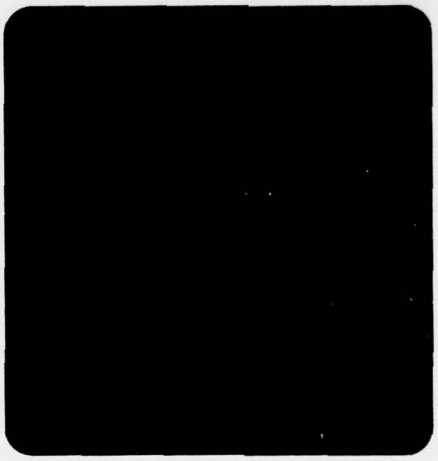
AD-A044 511 SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF INDUSTRIA--ETC F/G 9/2
METHODS FOR MODELLING AND GENERATING PROBABILISTIC COMPONENTS I--ETC(U)
AUG 77 B SCHMEISER N00014-77-C-0425
UNCLASSIFIED IEOR-77018 NL

1 OF 1
ADA044511

END
DATE FILMED
10-77
DDC

SOUTHERN METHODIST UNIVERSITY

AD A044511



DDC
SEP 23 1977
C

DEPARTMENT OF INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH
SCHOOL OF ENGINEERING AND APPLIED SCIENCE
DALLAS, TEXAS 75275

DDC FILE COPY
RJ NO.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

Technical Report IEOR-77018

METHODS FOR MODELLING AND GENERATING PROBABILISTIC COMPONENTS IN DIGITAL COMPUTER SIMULATION WHEN THE STANDARD DISTRIBUTIONS ARE NOT ADEQUATE: A SURVEY

Bruce Schmeiser

Department of Industrial Engineering
and Operations Research
Southern Methodist University
Dallas, Texas 75275

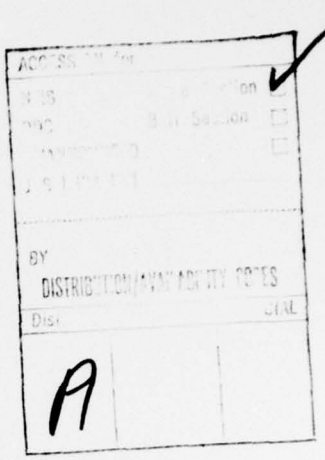
DDC
SEP 23 1977
C

August 1977

This work is supported in part by the Office of Naval Research under Contract N00014-77-C-0425.

This paper will be included in the Proceedings of the 1977 Winter Simulation Conference, Gaithersburg, MD, December 5-7.

409 434



ABSTRACT

Methods of modelling probabilistic components which are not adequately represented by the standard continuous distributions (such as normal, gamma, and Weibull) are surveyed. The methods are categorized as systems of distributions, approximations to the cumulative distribution function, and four-parameter distributions. Emphasis is on generality, determination of appropriate parameter values, and random variate generation.

I. INTRODUCTION

A common problem in computer simulation of stochastic systems is the selection of appropriate statistical distributions for generating random observations. The selection is sometimes an over-simplified process of selecting from a few well-known standard distributions: normal, lognormal, gamma, and Weibull (including the exponential) in the continuous case; Bernoulli, binomial, and Poisson in the discrete case. These distributions are appropriately chosen in many cases. Other times they are chosen because of the ease of their pseudo-random process generation.

Ideally, an appropriate statistical distribution is chosen using a three step iterative process consisting of

1. identification of a family of distributions which appear appropriate
2. determination of the parameter values which allow the family to best fit the situation, and
3. diagnostic checking to determine whether or not an adequate model has been found.

These three steps are repeated until the diagnostic check is satisfactory. (McGrath et al [28] give a detailed survey of techniques for selecting probability distributions.)

In those cases where the diagnostic check is not satisfactory for the standard distribution, more elaborate means must be used. This paper surveys methods useful for continuous random variables. In particular the following are discussed:

Systems of Distributions
- Pearson
- Johnson

Approximations of the inverse cdf
- Expansion techniques
- Polynomial regression
- Rectangular approximation

Four-parameter distributions
- Beta
- Four-parameter gamma
- Burr
- Generalized lambda
- Absolute lambda

The common property of these methods is their ability to satisfy four or more desired properties. The standard distributions usually satisfy one (exponential), two (normal, gamma, Weibull), or three (translated gamma or Weibull) properties. These properties may be moments, quantiles, location of the mode or other desired characteristics. The generality of a method depends on the number of properties satisfied and the combinations of properties possible.

The generality of an approach may be measured to some extent by considering the third and fourth standardized moments

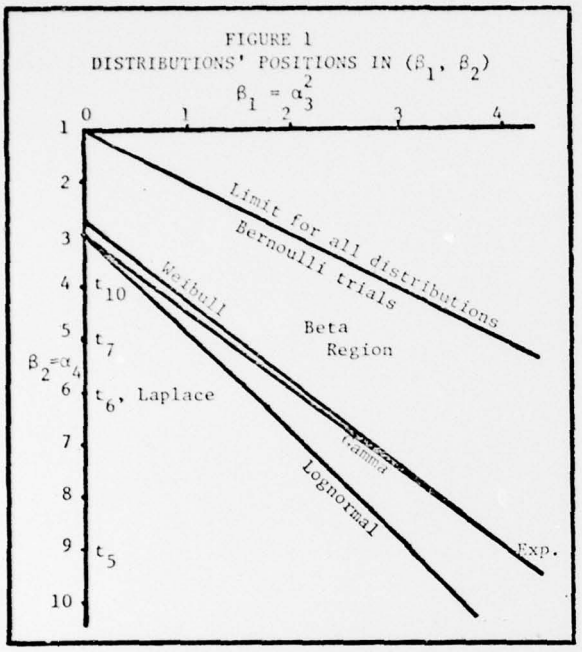
$$\alpha_i = E\{[(X-\mu)/\sigma]^i\} \qquad i = 3, 4.$$

where μ is the mean and σ the standard deviation of X. $\beta_1 = \alpha_3^2$ is a measure of skewness and the kurtosis (tail-weight and/or peakedness) may be measured by $\beta_2 = \alpha_4$. Since (β_1, β_2) is a measure of the shape of a distribution, the more values of (β_1, β_2) attainable by a method, the more general it is. While other measures are also reasonable, standardized moments will be used here to motivate these methods and to compare their generality.

Figure 1 shows the (β_1, β_2) plane. The locations of Bernoulli trials, uniform, normal, Weibull, gamma, exponential and Student-t are shown. Note that symmetric distributions lie on the $\beta_1 = 0$ axis. Bernoulli trials, the shortest tailed distribution possible, lie on the $\beta_2 - \beta_1 = 1$ line. The symmetric Bernoulli trial ($p = .5$) lies at the point $(\beta_1, \beta_2) = (0,1)$. The normal distribution lies at (0,3) and Student's t distribution lies below the normal. Asymmetric distributions are scattered around the square. The exponential is at $(\beta_1, \beta_2) = (4,9)$ which is a point on both the Weibull and gamma curves. Hahn and Shapiro [16] discuss the positions of the standard distributions in more

detail.

FIGURE 1
DISTRIBUTIONS' POSITIONS IN (β_1, β_2)



A characteristic of all of these distributions is that none covers a region of the (β_1, β_2) plane. The distributions usually considered general (Weibull, lognormal, gamma and Student's t) are associated with lines while more limited models (normal, exponential, uniform) are represented by points. It is clear that the standard distributions correspond to only a small part of the available models of randomness. While the standard techniques are often appropriate due to their theoretical foundation, when theoretical reasons do not justify a standard distribution, selection from any point on the plane is reasonable and often desirable.

The property common to all the versatile methods mentioned at the beginning of this section is that each covers a region of the (β_1, β_2) plane. To have this versatility, at least four parameters are necessary: two for shape (third and fourth moments), one for scaling (second moment -- variance) and one for location (first moment -- mean). Some models have more parameters.

The three groups of methods are examined in the next three sections. Generality, properties, techniques of parameter determination and techniques of random variate generation are discussed. The author apologizes for any omissions and would appreciate having them brought to his attention.

II. SYSTEMS OF DISTRIBUTIONS

The Pearson and Johnson systems each cover the entire (β_1, β_2) plane by using multiple functional forms. Each functional form covers a region of the (β_1, β_2) plane or is a transition between regions. Both systems were developed as statistical models before computer simulation was a factor.

PEARSON SYSTEM

The Pearson system [32] is composed of seven types (functional forms). Some correspond to standard distributions (beta, gamma, Student t) and the others provide a continuum over the (β_1, β_2) plane. There is exactly one Pearson distribution for each point of the (β_1, β_2) plane. The normal distribution is a limiting distribution of every type of Pearson distribution. Elderton and Johnson [13], Johnson and Kotz [24] and Ord [30] discuss the Pearson system in detail. Tadikamalla [36] discusses type selection, parameter estimation, and variate generation.

Due to the multiple forms, multiple parameter estimation and variate generation techniques are necessary to cover the whole (β_1, β_2) plane. Johnson, et al [23] give tables of percentage points of the Pearson curves for given $\sqrt{\beta_1}$ and β_2 which can aid parameter determination. Maximum likelihood estimation is difficult. However, parameters may be determined by matching the first four moments relatively easily. The Pearson system is one of the few approaches having closed form solution for matching moments. For example, Pearson types I and II (beta distribution) yield parameter values

$$p,q = r\{1 \pm (r+2)\ [\beta_1\{(r+2)^2\beta_1 + 16(r+1)\}^{-1}]^{1/2}\}/2$$

where

$$r = 6(\beta_2 - \beta_1 - 1)/(6 + 3\beta_1 - 2\beta_2)$$

and $p \lesseqgtr q$ according to whether $\sqrt{\beta_1} \lesseqgtr 0$. The only drawback is that the equations change for each type. However, the type may be found easily from (β_1, β_2) charts usually reproduced whenever the Pearson system is discussed.

Variate generation is not straightforward. Each type requires its own techniques. The inverse cdf is not closed form for the more important types and rejection techniques have not worked well until recently due to many types having long tails. Cooper, Davis, and Dono [9] describe their PURGE program which calculates moments and generates values using numerical integration. Numerical integration was important at that time since no exact techniques were known. Since then much development has occurred in generation techniques for the Pearson system. In particular, type I and II Pearson (beta) variate generation has been the subject for Jöhnk [19], Fishman [15], Whittaker [41], Ahrens and Dieter [1], Atkinson and Pearce [2], Atkinson and Whittaker [3], and Schmeiser and Shalaby [35]. These developments are important to the whole Pearson family since beta variates can be transtormed to obtain other types difficult to generate directly. Using the beta relationship, McGrath and Irving [29] give flow charts and FORTRAN subroutines for each Pearson type.

JOHNSON SYSTEMS

Johnson [20, 21] has developed two systems which cover the entire (β_1, β_2) plane in three functional forms based on transformations of the normal and Laplace distributions. Let z be a standard normal variate. Then

$$x = \xi + \exp[(z-\gamma)/\delta] \qquad \xi \le x$$

$$x = \frac{\xi + (\xi+\lambda)\exp[(z-\gamma)/\delta]}{1 + \exp[(z-\gamma)/\delta]} \qquad \xi \le x \le \xi + \lambda$$

and

$$x = \xi + \lambda \sinh[(z-\gamma)/\delta] \qquad -\infty < x < \infty$$

correspond to the three types S_L (lognormal), S_B (bounded) and S_U (unbounded), yielding straightforward, though relatively slow, variate generation. The second system is obtained by letting z be a standard Laplace random variate. For both systems ξ is a location parameter, λ is a scale parameter, and γ and δ are shape parameters.

Parameter estimation can proceed in at least two ways: matching percentiles or matching moments. Johnson [22] gives tables for γ and δ in terms of $\sqrt{\beta_1}$ and β_2 for the S_U family. Johnson and Kotz [24] give an abac with the same information in graphical form. Given γ and δ, the desired variance σ^2 is obtained from

$$\lambda = \sigma/\sigma(y)$$

and the desired mean μ is obtained from

$$\xi = \mu - \lambda E(y)$$

where

$$E(y) = -\omega^{1/2} \sinh \Omega$$

$$\sigma(y) = [(\omega-1)(\omega \cosh 2\Omega+1)/2]^{1/2}$$

$$\omega = \exp(\delta^{-2})$$

and

$$\Omega = \gamma/\delta$$

For the S_B family, ξ is the lower bound and λ is the range, reducing the problem to only two parameters γ and δ. Given any two desired percentiles (x_1, p_1) and (x_2, p_2) yields

$$\delta = \frac{z_1 - z_2}{\ln\left[\frac{x_1-\xi}{\xi+\lambda-x_1}\right] - \ln\left[\frac{x_2-\xi}{\xi+\lambda-x_1}\right]}$$

and

$$\gamma = z_1 - \delta \ln\left[\frac{x_1-\xi}{\xi+\lambda-x_1}\right]$$

where

$z_i = \Phi(p_i)$, the p_i^{th} percentile of the standard normal distribution.

Similarly for the S_L distribution, ξ is the lower bound and

$$\delta = \frac{z_1 - z_2}{\ln(x_1-\xi) - \ln(x_2-\xi)}$$

and

$$\gamma = z_1 - \ln(x_1-\xi)$$

yields two desired percentiles. Tadikamalla [36] also discusses these topics.

III. APPROXIMATIONS TO THE INVERSE CDF

Three techniques based on approximating the inverse cumulative distribution function (cdf) have been proposed: expansion techniques, polynomial regression, and rectangular approximation. Although all three only approximate the desired cdf, the approximation can be made as accurate as desired at the expense of coding, storage and/or executive time.

EXPANSION TECHNIQUES

Expansion techniques can sometimes provide a random variable having four or more desired moments. The general technique is to expand the density or cdf into a series of infinite terms based on some initial distribution which, if used in its entirety, would provide the desired moments exactly. The initial distribution is usually the normal, although others such as the gamma have been used. Elderton and Johnson [13], Johnson and Kotz [24] and Ord [30] discuss expansion techniques in more detail. Only the special case of the Cornish-Fisher expansion [10] is discussed here.

The Cornish-Fisher expansion, which uses a normal initial distribution, can be used to generate random variates having desired moments μ_1' (mean), μ_2 (variance), μ_3 (third central moment), ..., μ_k (k^{th} central moment). We consider $k = 4$ here. Set

$$\kappa_1 = \mu_1'$$

$$\kappa_2 = \mu_2$$

$$\kappa_3 = \mu_3$$

$$\kappa_4 = \mu_4 - \mu_2^2$$

which are the first four cumulants, upon which the Cornish-Fisher expansion is based. Now generate a normal (0,1) variate z. Setting

$$\begin{aligned} x = z &+ (z^2-1)\kappa_3/6 + (z^3-3z)\kappa_4/24 \\ &-(2z^3-5z)\kappa_3^2/36 - (z^4-5z^2+2)\kappa_3\kappa_4/24 \\ &+ (12z^4-53z^2+17)\kappa_3^3/324 \\ &- (3z^5-24z^3+29z)\kappa_3^2\kappa_4/384 \\ &+ (14z^5-103z^3+107z)\kappa_3^2\kappa_4/288 \\ &- (252z^5-1688z^3+1511z)\kappa_3^4/7776 \\ &+ \ldots \end{aligned}$$

yields $v = \kappa_1 + \sqrt{\kappa_2}\, x$ as the desired variate. Fisher and Cornish [14] give terms through the eighth

cumulant.

Expansion techniques converge rapidly when the distribution desired is similar to the initial distribution, but later terms become important otherwise. (Barton and Dennis [5] discuss the region of (β_1, β_2) where the resulting distribution is unimodal and has all nonnegative density.) Therefore care must be taken when using this technique. A simple procedure is to collect sample moments on the generated observations to ensure the desired moments are obtained.

POLYNOMIAL REGRESSION

Tocher [38] uses an approximation to the inverse cdf

$$F^{-1}(u) = a+bu+cu^2+\alpha(1-u)^2\log(u)+\beta u^2\log(1-u)$$

where a, b, c, α, and β are constants used to fit the desired distribution. Variate generation is trivial by setting $x = F^{-1}(u)$ where u is a uniform (0,1) observation. While X will have some distribution, $F^{-1}(u)$ is not an inverse cdf for all parameter values. Also the fit may or may be accurate enough for the intended purpose. Tocher notes that if five particular quantiles are desired, (x_i, u_i) i = 1, 2, 3, 4, 5, then the resulting five equations (from substitution into the inverse cdf) are linear and can therefore be solved relatively easily. In general, he suggests least squares estimation to satisfy desired properties as nearly as possible.

RECTANGULAR APPROXIMATION

Rectangular approximation is probably the most widely used technique for generating random variates which do not come from the standard distributions. (GPSS, for example, uses this technique.) The (inverse) cdf is approximated by a piecewise linear function, which is equivalent to a mixture of k nonoverlapping adjacent uniform distributions. To generate random variates, a uniform value is generated from the i^{th} distribution with probability p_i, i = 1, 2, ..., k. In the special case of $p_i = 1/k$ for all i, storage is reduced since the p_i's no longer need to be stored. McGrath and Irving [29] and Barnard and Cawdery [4] point out that the constant p_i values case can also be made faster than general rectangular approximation when k is large.

While variate generation is direct, the proper piecewise linear approximation is not so obvious. Drake and Gerhardt [11] and McGilchrist [27] study this problem in a general context. Kisco [26] discusses this problem in the simulation context, presenting a 27 step algorithm to fit empirical data. Even with an algorithmic approach it is difficult to obtain the desired moments. For example, in Kisco's example for the algorithmic approach, the mean is .832 units too large. This may or may not be important in a given situation. Of course a nonlinear programming problem could be solved along the lines of minimizing average error subject to satisfying the desired mean and variance. But as the number of linear segments grows, and therefore the error decreases, the difficulty in obtaining desired moments increases. Nevertheless, this technique works even when the desired distribution is not unimodal, which is not true of the other methods discussed here. Butler [8] extends rectangular approximation to the equivalent of quadratic approximation by using rejection methods on a piecewise linear density function.

IV. FOUR-PARAMETER DISTRIBUTIONS

There are five distributions which cover wide regions of the (β_1, β_2) plane: beta (already discussed under the Pearson system), four parameter gamma, Burr, generalized lambda distribution, and a distribution referred to here as the absolute lambda distribution. All have exactly four parameters and only one functional form, resulting in parameter estimation and variate generation methods which are applicable in many modelling situations where the standard distributions are not adequate.

FOUR-PARAMETER GAMMA DISTRIBUTION

Harter [17] discusses a four-parameter distribution which includes several standard distributions as special cases. The density function is

$$f(x;c,a,b,p)=p(x-c)^{bp-1}\exp\{-[(x-c)/a]^p\}/a^{bp}\Gamma(b)$$

where $a,b,p > 0$ and $x \geq c$. Here c is the location parameter, a is the scale parameter and b and p together determine shape. When c = 0, this distribution reduces to the gamma if p = 1, Weibull if b = 1, exponential if b = p = 1, and half-normal if b = 1/2, p=2. In addition, the normal distribution is a limiting case as p = 1 and b goes to infinity.

While being more general than either the Weibull or gamma distribution, this distribution has several disadvantages for use in simulation. First, the moments are not easy to evaluate except for the special cases. Second, parameter estimation is not straightforward. Harter [17] discusses maximum likelihood estimation including the estimators' asymptotic variances and covariances. The third disadvantage is in random variate generation. Since the standard gamma distributions (Pearson type III) are not straightforward to generate (except for integer shape parameter), certainly this four parameter generalization is difficult to generate. The author knows of no results in this area.

BURR DISTRIBUTION

The Burr distribution has cdf

$$F(x) = 1 - \{1 + [(x-a)/b]^c\}^{-k} \qquad x \geq 0$$

with parameters b and k positive. Here a is the location parameter, b the scale parameter, and c and k determine the shape. As originally developed by Burr [6], the parameter c was positive. As shown by Wheeler [40] the region of (β_1, β_2) covered

by this distribution includes the Weibull and gamma curves and some area corresponding to heavier tails (larger β_2). Burr [7] notes that using $c<0$ results in coverage of a new region of (β_1, β_2) closer to the line $\beta_2 - \beta_1 = 1$ in the beta region. Wheeler [40] and Tadikamalla and Ramberg [37] have used the Burr distribution to approximate the gamma distribution.

The parameters may be estimated using tables given in Burr [7] to match the first four moments. Maximum likelihood estimators can be obtained via iterative numerically techniques. Likewise trying to fit the parameters by specifying quantiles requires numerical solution. Random variate generation is direct from the inverse cdf

$$x = F^{-1}(u) = a + b\,[(1-u)^{-1/k}-1]^{1/c}$$

where u is a uniform (0,1) variate.

GENERALIZED LAMBDA DISTRIBUTION

Tukey [39] considered the inverse cdf

$$x = F^{-1}(u) = [u^{\lambda} - (1-u)^{\lambda}]/\lambda \qquad (0 \le u \le 1)$$

with parameter $\lambda > 0$. This distribution is symmetric about zero and has a wide range $(1.75,\infty)$ of values of β_2. Joiner and Rosenblatt [25] noted that for $\lambda \doteq .14$ the distribution closely approximated the normal distribution.

Ramberg and Schmeiser [32, 33] generalized the distribution to

$$x = F^{-1}(u) = \lambda_1 + [u^{\lambda_3} - (1-u)^{\lambda_4}]/\lambda_2 \qquad (0 \le u \le 1)$$

where λ_2, λ_3, λ_4 have the same sign. The location parameter is λ_1, the scale parameter is λ_2, and shape is determined by λ_3 and λ_4. This distribution covers all of the (β_1, β_2) plane except that corresponding to the U - shaped beta distribution. Combined with the U - shaped beta, the entire (β_1, β_2) plane may be covered in two functional forms by noting that the limiting distribution of the generalized lambda distribution as either λ_3 or λ_4 goes to infinity is the power function distribution, which is a special case of the beta family for one parameter equal to one.

Nearly exponential and nearly normal variates may be generated from the generalized lambda distribution. The exponential is a limiting distribution as $\lambda_1 = \lambda/\lambda_3$, $\lambda_2 = \lambda_2/\lambda$, λ_3 goes to infinity and λ_4 goes to zero. The standardized normal distribution is approximated by the parameter values $\lambda_1 = 0$, $\lambda_2 = .1975$ and $\lambda_3 = \lambda_4 = .1349$, which match the first four moments. The maximum deviation between the two cdf's is about .001. (This is about one-tenth the maximum error obtained using the central limit theorem approximation with twelve uniform variates.) While not fast (Atkinson and Pearce [2]), this does yield a one-line normal generator for situations where such an approximation is reasonable. (There are many faster and exact techniques for normal generation, which you probably already know if you have read this far.)

The close approximation to the normal is important when sensitivity analysis is to be performed on normally distributed values. Using this one functional form, Hogg, Fisher, and Randles [18], for example, use the generalized lambda distribution to test the behavior of a distribution-free test statistic on both normal and non-normal random variables.

Dudewicz, Ramberg, and Tadikamalla [12] give tables to determine the parameters to match the first four moments. Tadikamalla [36] discusses fitting any four percentiles exactly and using least squares to fit the distribution to more than four percentiles, both of which must be done using iterative methods. Random variate generation is straightforward using the inverse cdf.

ABSOLUTE LAMBDA DISTRIBUTION

Schmeiser and Deutsch [34] discuss

$$x = F^{-1}(u) = \lambda_1 + \lambda_2\,[g(\lambda_3, \lambda_4)] \quad (0 \le u \le 1)$$

where $g(\lambda_3, \lambda_4) = \begin{cases} -(\lambda_4 - u)^{\lambda_3} & \text{if } u \le \lambda_4 \\ (u - \lambda_4)^{\lambda_3} & \text{if } u > \lambda_4 \end{cases}$

where $-\infty < \lambda_1 < \infty$, $0 < \lambda_2$, $0 \le \lambda_3$, and $0 \le \lambda_4 \le 1$ which is a mixture of two distributions. To the author's knowledge, this is the only distribution having a single functional form which covers the entire (β_1, β_2) plane. Three distributions are obtained exactly: Bernoulli trials if $\lambda_3 = 0$, uniform if $\lambda_3 = 1$, and exponential as λ_3 goes to zero with $\lambda_1 = \lambda_2 = \lambda/\lambda_3$ and $\lambda_4 = 1$.

Parameters may be determined using only closed form expressions by matching percentiles. Given the mode, Prob $\{x \le \text{mode}\}$, and any two percentiles (x_1, p_1) and (x_2, p_2), the corresponding parameters are $\lambda_1 = \text{mode}$, $\lambda_4 = \text{Prob}\{x \le \text{mode}\}$,

$$\lambda_3 = \frac{\ln|(\lambda_1 - x_1)/(\lambda_1 - x_2)|}{\ln|(\lambda_4 - p_1)/(\lambda_4 - p_2)|}$$

and

$$\lambda_2 = |\lambda_1 - x_1|/|\lambda_4 - p_1|^{\lambda_3}$$

Moments may be matched using graphs in Schmeiser and Deutsch [34]. Random number generation is again straightforward due to the closed form inverse cdf.

An important disadvantage of the absolute lambda distribution is its truncated tails and extreme value of the density function at the mode. In particular, the density function at $x = \lambda_1$, takes on only the values 0, 1, or infinity as $\lambda_3 < 1$, $\lambda_3 = 1$, or $\lambda_3 > \infty$, respectively. This behavior is similar to that of the gamma distribution as the shape parameter $\alpha > 1$, $\alpha = 1$, or $\alpha < 1$. However it is a more serious drawback at the mode than at an end point. Current research is oriented toward replacing the center of the distribution with a "nice" function such as a quadratic.

V. CONCLUSION

Available to the simulation practitioner are many techniques which allow the modelling and generation of probabilistic components when the standard distributions are not adequate. The techniques vary in generality, ease of parameter estimation, ease of variate generation, and speed of execution. The technique selected should reflect the particular application at hand.

REFERENCES

1. Ahrens, J. H. and Dieter, U., "Computer methods for sampling from gamma, beta, Poisson, and binomial distributions," Computing, 12(1974), 223-246.
2. Atkinson, A. C. and Pearce, M. C., "The computer generation of beta, gamma, and normal random variables," Journal of the Royal Statistical Society, A, 139(1976), 431-461.
3. Atkinson, A. C. and Whittaker, J., "A switching algorithm for the generation of beta random variables with at least one parameter less than one," Journal of the Royal Statistical Society, A, 139 (1976), 462-467.
4. Barnard, D. R. and Cawdery, M. N., "A note on a new method of histogram sampling," Operational Research Quarterly, 25(1974), 319-320.
5. Barton, D. E. and Dennis, K. E. R., "The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal," Biometrika, 39(1952), 425-427.
6. Burr, I. W., "Cumulative frequency functions," Annals of Mathematical Statistics, 13(1942), 215-232.
7. Burr, I. W., "Parameters for a general system of distributions to match a grid of α_3 and α_4," Communications in Statistics, 2(1973), 1-21.
8. Butler, E. G., "Algorithm 370, general random number generator," Communications of the ACM, 13 (1970), 49-51.
9. Cooper, J. D., Davis, S. A., Dono, N. R., "Pearson Universal Random Distribution Generator (PURGE)," 1965 ASQC Transactions, 402-411.
10. Cornish, E. A. and Fisher, R. A., "Moments and cumulants in the specification of distributions," Institut International de Statistique, Revue, 4 (1937), 1-14.
11. Drake, K. W. and Gerhardt, L. A., "A class of pdf modeling algorithms," IEEE Transactions on Systems, Man, and Cybernetics, 2(1972), 402-407.
12. Dudewicz, E. J., Ramberg, J. S. and Tadikamalla, P. R., "A distribution for data fitting and simulation," 1974 ASQC Transactions, 28, 407-418.
13. Elderton, W. P. and Johnson, N. L., Systems of Frequency Curves, Cambridge University Press, Cambridge, 1969.
14. Fisher, R. A. and Cornish, E. A., "The percentile points of distributions having known cumulants", Technometrics, 2(1960), 209-225.
15. Fishman, G. S., Concepts and Methods in Discrete Event Digital Simulation, Wiley, New York, 1973.
16. Hahn, G. J. and Shapiro, S. S., Statistical Models in Engineering, Wiley, New York, 1967.
17. Harter, H. L., "Maximum likelihood estimation of the parameters of a four-parameter generalized gamma population from complete and censored samples," Technometrics, 9(1967), 159-165.
18. Hogg, R. V., Fisher, D. M., and Randles, R. H., "A two-sample adaptive distribution-free test," Journal of the American Statistical Association, 70(1975), 656-661.
19. Johnk, M. D., "Erzeugung von Betaverteilten und Gammaverteilten Zufallszahlen," Metrika, 8(1964), 5-15.
20. Johnson, N. L., "Systems of frequency curves generated by methods of translation," Biometrika, 36(1949), 149-176.
21. Johnson, N. L., "Systems of frequency curves derived from the first law of Laplace," Trabajos de Estadistica, 5(1954), 283-291.
22. Johnson, N. L., "Tables to facilitate fitting S_U frequency curves," Biometrika, 52(1965), 547-558.
23. Johnson, N. L., Nixon, E., Amos, D. E. and Pearson, E. S., "Table of percentage points of Pearson curves, for given $\sqrt{\beta_1}$ and β_2, expressed in standard measure," Biometrika, 50(1963), 459-98.
24. Johnson, N. L., Kotz, S. Distributions in Statistics: Continuous Univariate Distributions -1, Wiley, New York, 1970.
25. Joiner, B. L. and Rosenblatt, J. R., "Some properties of the range in samples from Tukey's symmetric lambda distribution," Journal of the American Statistical Association, 66(1971), 394-399.
26. Kisco, T. M., "An automated method of creating piecewise linear cumulative probability distributions," 1976 Winter Simulation Conference Proceedings, 487-494.
27. McGilchrist, C. A., "Estimation of unimodal probability," Sankhyā, 37A (1975), 139-149.
28. McGrath, E. J., Basin, S. L., Burton, R. W., Irving, D. C., Jaquette, S. C., Ketler, W. R., and Smith, C. A., Techniques for Efficient Monte Carlo Simulation, Volume I: Selecting Probability Distributions, NTIS, Springfield, Va., 1973.
29. McGrath, E. J. and Irving, D. C., Techniques for Efficient Monte Carlo Simulation, Volume II: Random Number Generation for Selected Probability Distributions, NTIS, Springfield, Va., 1973.
30. Ord, J. K., Families of Frequency Distributions, Hafner, New York, 1972.
31. Pearson, K., "Contributions to the mathematical theory of evolution. II. Skew variations in homogeneous material," Philosophical Transactions of the Royal Society of London, Series A, 186(1895), 343-414.
32. Ramberg, J. S. and Schmeiser, B. W., "An approximate method for generating symmetric random variables," Communications of the ACM, 15(1972), 987-990.
33. Ramberg, J. S. and Schmeiser, B. W., "An approximate method for generating asymmetric random variables," Communications of the ACM, 17(1974), 78-82.
34. Schmeiser, B. W. and Deutsch, S. J., "A versatile four parameter family of probability distributions suitable for simulation," AIIE Transactions, 9(1977), 176-182.
35. Schmeiser, B. W. and Shalaby, M., "Rejection methods for beta variate generation," Technical Report IEOR 77014, Southern Methodist University, July 1977.
36. Tadikamalla, P. R., Modeling and Generating Stochastic Inputs for Simulation Studies, Ph.D. Dis-

sertation, The University of Iowa, 1975.
37. Tadikamalla, P. R. and Ramberg, J. S., "An approximate method for generating gamma and other variates," Journal of Statistical Computation and Simulation, 3(1975), 275-282.
38. Tocher, K. D., The Art of Simulation, English Universities Press, London, 1963.
39. Tukey, J. W., Technical Report 36, Statistical Techniques Research Group, Princeton University, 1960.
40. Wheeler, D. J., "An approximation for simulation of gamma distributions," Journal of Statistical Computation and Simulation, 3(1975), 225-232.
41. Whittaker, J., "Generating gamma and beta random variables with non-integral shape parameter," Applied Statistics, 23(1974), 210-214.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER IEOR 77018	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Methods for Modelling and Generating Probabilistic Components in Digital Computer Simulation When the Standard Distributions are Not Adequate: A Survey		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Bruce Schmeiser		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0425
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Dallas, TX 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Analysis Program (Code 431) Office of Naval Research Arlington, VA 22217		12. REPORT DATE August 1977 13. NUMBER OF PAGES 7
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Simulation
Random Variates
Monte Carlo
Process Generation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Methods of modelling probabilistic components which are not adequately represented by the standard continuous distributions (such as normal, gamma, and Weibull) are surveyed. The methods are categorized as systems of distributions, approximations to the cumulative distribution function, and four-parameter distributions. Emphasis is on generality, determination of appropriate parameter values, and random variate generation.